

OIIPE

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/974,546

DATE: 12/18/2001

TIME: 16:40:03

Input Set : N:\Crf3\RULE60\09974546.raw

Output Set: N:\CRF3\12182001\I974546.raw

## SEQUENCE LISTING

## 3 (1) GENERAL INFORMATION:

## 5 (i) APPLICANT: An, Gang

6 O'Hara, S. Mark

7 Ralph, David

8 Veltri, Robert

## 10 (ii) TITLE OF INVENTION: BIOMARKERS AND TARGETS FOR DIAGNOSIS,

11 PROGNOSIS AND MANAGEMENT OF PROSTATE DISEASE

## 13 (iii) NUMBER OF SEQUENCES: 87

## 15 (iv) CORRESPONDENCE ADDRESS:

16 (A) ADDRESSEE: Arnold, White &amp; Durkee

17 (B) STREET: P.O. Box 4433

18 (C) CITY: Houston

19 (D) STATE: Texas

20 (E) COUNTRY: USA

21 (F) ZIP: 77210

## 23 (v) COMPUTER READABLE FORM:

24 (A) MEDIUM TYPE: Floppy disk

25 (B) COMPUTER: IBM PC compatible

26 (C) OPERATING SYSTEM: PC-DOS/MS-DOS

27 (D) SOFTWARE: PatentIn Release #1.0, Version #1.30

## 29 (vi) CURRENT APPLICATION DATA:

C--&gt; 30 (A) APPLICATION NUMBER: US/09/974,546

C--&gt; 31 (B) FILING DATE: 10-Oct-2001

32 (C) CLASSIFICATION:

## 34 (vii) PRIOR APPLICATION DATA:

35 (A) APPLICATION NUMBER: 09/097,199

36 (B) FILING DATE:

## 38 (viii) ATTORNEY/AGENT INFORMATION:

39 (A) NAME: Nakashima, Richard A.

40 (B) REGISTRATION NUMBER: P-42,023

41 (C) REFERENCE/DOCKET NUMBER: UROC:018

## 43 (ix) TELECOMMUNICATION INFORMATION:

44 (A) TELEPHONE: (512) 418-3000

45 (B) TELEFAX: (512) 474-7577

## 48 (2) INFORMATION FOR SEQ ID NO: 1:

## 50 (i) SEQUENCE CHARACTERISTICS:

51 (A) LENGTH: 391 base pairs

52 (B) TYPE: nucleic acid

53 (C) STRANDEDNESS: single

54 (D) TOPOLOGY: linear

## 56 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

58	GTCCAGTCGC TCAGAAATTT CCTTTGATGC TTTGAAGTTA TCTCTCTTGG ATCTGCTTCC	60
60	TCCTTATCGT CTCTACATCC CAAGAACAGA GAGTGAGTCT TCTTTATTTT CTTATCTCTG	120
62	TTTTTAGCAC AGTATTTGAT ATATAGTGTA GATACTATAA ATGCTTGCTA AACTTTGTCA	180
64	AATTCCACAT TTTTAAATA AAAATGAGAA TGAGCTTGTA GTCAACATGG CGTTTGTAAG	240
66	TTTGGAGTCT ATATATGGTA GATATACATA TTTTAAATC TAAGTGCAAC TTTTCTCTTG	300

ENTERED

1601  
1635  
#4 / K.T.  
1/9  
Raw RECEIVED  
Set  
12/18/2001  
TECH CENTER 1600/2900

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/974,546

DATE: 12/18/2001

TIME: 16:40:03

Input Set : N:\Crif3\RULE60\09974546.raw

Output Set: N:\CRF3\12182001\I974546.raw

68 ATTATCTTGA AATGCCTTAT CATCTCCACA TTTGCTGTAG GCAGTAGTTT AGTGGGTCCA 360  
70 TTATATCTGC CACACTGATT GTCTTAAATA A 391  
73 (2) INFORMATION FOR SEQ ID NO: 2:  
75 (i) SEQUENCE CHARACTERISTICS:  
76 (A) LENGTH: 614 base pairs  
77 (B) TYPE: nucleic acid  
78 (C) STRANDEDNESS: single  
79 (D) TOPOLOGY: linear  
81 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
83 CAGTAGTGGC CCCAAATGCC AGGCTGCACT GATATTTATT GGATATAAGA CAAAGGGGCA 60  
85 GGGTAAGGAA TGTGAACCAT CTCCAATAAT AGGTAAGGTC ACATGGGTCA TGTGTCCACT 120  
87 GGACAGGGGG CCCTTCCCTG CCTGGCAGCA GAGGCAGAGA GAGAGAGAAG AGAGAGAGAC 180  
89 AGCTTATGCC ATTATTTCTG CATATCAGAC ATTTAGTACT TTCACTAATT TGCTCCTGCT 240  
91 ATCTAAAAGG CAGAGCCAGG TATACAGGAT GGAACATGAA AGCGGACTAG GAGCGTGACC 300  
93 ACTGAAGCAC AGCATCACAG GGAGACAGGC CTCTGGATAC TGGCCGGGGG GCCCTGACTG 360  
95 ATGTCAAGGC CCTCCACAAG AGTGGAGGAG TTAGTCTTCC TCTAAACTCC CCCGGGGGAA 420  
97 AGGGAGGCTC CTTTTCCCAG TCTGCTAAGT AGTGGGTGTT TTTCTTGAC ACTGATGCTA 480  
99 CTGCTAGACC ATGGTCCACT TTGCAACAGG CATCTTCCCA GACACTGGTG TTAGTGCTAG 540  
101 ACCAAGCCCT CTGGTGGCCC TGTCCGGGCA TAAGAGAAGG CTCACACTCT TGTCTTCTGG 600  
103 CCACTTCGCA CTAT 614  
106 (2) INFORMATION FOR SEQ ID NO: 3:  
108 (i) SEQUENCE CHARACTERISTICS:  
109 (A) LENGTH: 757 base pairs  
110 (B) TYPE: nucleic acid  
111 (C) STRANDEDNESS: single  
112 (D) TOPOLOGY: linear  
114 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:  
116 ACAACGACAC ATTCAAGAGT TAAATATTTA TCATCAAACA TTGGATTTTT CCTTAACGCT 60  
118 AGAGATTGCT ACAAACTTTC TGAAGGGTCT CAATGGCTTC AGGCTAAGAA GAGATTTCTC 120  
120 CCTGTATATA GCAGCAAGAC AAATTAGCCA TTTCACTCTC AAACCTTCACT AATGATCACA 180  
122 TTCTTTCCAA AAGGAACCTCT AGAAGACCAA ATGCCCCGAG TTAAGAACAT CAAAACTAAC 240  
124 CATCTGAAGA AACTTCCCAA GTGTAAGACT CTGCCATTAA AACATTACCG AGAGGGGACT 300  
126 CAAACAGTCT TTTCTTCCCT TTGTCTGTGT TCTTTGCTCC CAGACCCAAG GCACTTGGCG 360  
128 GACAGTACTT GATACAATAA TTTAAAAAGC ACCACTCCCT TCCCACCTTG TAAATACCCA 420  
130 GAACTCTAAT TGGACCACCC TGAAGCTTAG GACCTACCAG CCATACAAAT AGTAAACTCT 480  
132 GTCCACGATT CACTCATCTG TGTATTTTCT ATAGATGTTT ACTAGGCGTT TGTTATATAA 540  
134 AAATACCCCG GCCAGGCACG GTGGCTCACG CCTGTAATCC CAGCACTTTG GGAGGTGGGT 600  
136 GGATCACCTG AGGTCTGGGAG TTCGAGACCA GCCTGACCAG CATGGTGGAA CCCCACCTC 660  
138 TACTAAAAAC ACAAAAATT AGCCGGGCGT GGTGGCACAT GCCTGTAATC CCAGCTACTC 720  
140 AGGAGGCTGA GCGGAGAAT TGCTTGAACC CGGAAGG 757  
143 (2) INFORMATION FOR SEQ ID NO: 4:  
145 (i) SEQUENCE CHARACTERISTICS:  
146 (A) LENGTH: 673 base pairs  
147 (B) TYPE: nucleic acid  
148 (C) STRANDEDNESS: single  
149 (D) TOPOLOGY: linear  
151 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:  
153 CAGGACACAG AGTAAGATAC CCACTGACTT CTTGTGGTCT ACTTCCTGGG TGTTGTTTCA 60  
155 ATGGGCTTTG TTATAACAGG ACTAGTCTTC TGTAATACA ACTTGGTAAA TAGGATGAAA 120

RECEIVED  
JAN 08 2002  
TECH CENTER 1600/2900

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/974,546

DATE: 12/18/2001

TIME: 16:40:03

Input Set : N:\Crif3\RULE60\09974546.raw

Output Set: N:\CRF3\12182001\I974546.raw

```

157 CATAACTTTG CGACAATTCA GTAGAAATAG GCATACAAAC CTGGGCCTGA TGACACTCAC      180
159 CTCCCCTTGG CTATAAACAT TACCCTACCT GTTAAGTCAG TAATCCTTTG GGAGAGCGCT      240
161 TACTGAGTAT CTATGATATG CAAAGACCAA AGACCGAGGG GGATCCCTGG TGTAGAGCAA      300
163 GCACACACCT GGTTATTAGC TACCTGCCAC CCTGCTGGGC ATGCAACATA CATTGTCTCA      360
165 AATTCTAACC ACCCTGCAAG GCAAGCTTCC TTGTTCTTTT AAAGAAGAAA AGTAGACCAG      420
167 CAAGATTGAT TTGCTCAAGA TTACACAGCC TGAATCTTG TCATGGGCAT GTCTGACTCT      480
169 GATAGCAATA CCCTCAAAGA AACTGTCAGA GAAGACTCAA TAAGAAGAAA GTTGAGATAC      540
171 AGAAACCAAC AGGAGAAGGT AATTCAGAAA TTCAAACAGA GTGGGTGTGA TGGGAAGAAT      600
173 TCATTAATAA GAAGGTACCT CTGTAGAAAA ATCTTACCAG ACAGTCTGGA AGTGAAGGAA      660
175 ACAGCCAATA GTC                                                    673
178 (2) INFORMATION FOR SEQ ID NO: 5:
180     (i) SEQUENCE CHARACTERISTICS:
181         (A) LENGTH: 358 base pairs
182         (B) TYPE: nucleic acid
183         (C) STRANDEDNESS: single
184         (D) TOPOLOGY: linear
186     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:
188 GTCACTGCAC ATTAAGATGG AGCCGAAGA GCCACACTCC GAGGGGGCAT CGCAGGAGGA      60
190 TGGGGCTCAA GGTGCCTGGG GCTGGGCACC CCTAAGTCAC GGCTCTAAGG AGAAAGCTCT      120
192 CTTCTGCCC GCGGAGCCC TCCCCTCCCC CCGGATCCCC GTGCTTTCCC GAGAGGGGAG      180
194 GACCAGAGAC CGGCAGATGG CTGCAGCGCT CCTCACTGCC TGGTCCCAGA TGCCAGTGAC      240
196 TTTTCGAGGAT GTGGCCTTGT ACCTCTCCCG GGAGGAGTGG GGACGGCTGG ACCACACGCA      300
198 GCAGAACTTC TACAGGGAAT GTCCTGCAGA AGAAAAATGG GCTGTCACTG GGCTTTCC      358
201 (2) INFORMATION FOR SEQ ID NO: 6:
203     (i) SEQUENCE CHARACTERISTICS:
204         (A) LENGTH: 23 base pairs
205         (B) TYPE: nucleic acid
206         (C) STRANDEDNESS: single
207         (D) TOPOLOGY: linear
209     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:
211 CACAGATGTA GCTTCCTCAC TGG                                                    23
214 (2) INFORMATION FOR SEQ ID NO: 7:
216     (i) SEQUENCE CHARACTERISTICS:
217         (A) LENGTH: 610 base pairs
218         (B) TYPE: nucleic acid
219         (C) STRANDEDNESS: single
220         (D) TOPOLOGY: linear
222     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:
224 CTGGAGTACA ATGTCAGTGT TTACTCTGTC AAGGATGACA AGGAAAGTGT CCCTATCTCT      60
226 GATACCATCA TCCCAGCTGT TCCTCCTCCC ACTGACCTGC GATTACCAA CATTGGTCCA      120
228 GACACCATGC GTGTCACCTG GGCTCCACCC CCATCCATTG ATTTAACCAA CTTCTGCTG      180
230 CGTTACTCAC CTGTGAAAAA TGAGGAAGAT GTTGACAGAGT TGTCAATTTT TCCTTCAGAC      240
232 AATGCAGTGG TCTTAACAAA TCTCCTGCCT GGTACAGAAT ATGTAGTGAG TGTCTCCAGT      300
234 GTCTACGAAC AACATGAGAG CACACCTCTT AGAGGAAGAC AGAAAACAGG TCTTGATTCC      360
236 CCAACTGGCA TTGACTTTTC TGATATTACT GCCAACTCTT TTAAGTGCA CTGGATTGCT      420
238 CCTCGAGCCA CCATCACTGG CTACAGGATC CGCCATCATC CCGAGCACTT CAGTGGGAGA      480
240 CCTCGAGAAG ATCGGGTGCC CCACTCTCGG AATTCCATCA CCCTCACCAA CCTCACTCCA      540
242 GGCACAGAGT ATGTGGTCAG CATCGTTGCT CTTAATGGCA GAGAGGAAAG TCCCTTATTG      600
244 ATTGGCCAAC                                                    610

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/974,546

DATE: 12/18/2001

TIME: 16:40:03

Input Set : N:\Crif3\RULE60\09974546.raw

Output Set: N:\CRF3\12182001\I974546.raw

## 247 (2) INFORMATION FOR SEQ ID NO: 8:

## 249 (i) SEQUENCE CHARACTERISTICS:

250 (A) LENGTH: 1649 base pairs

251 (B) TYPE: nucleic acid

252 (C) STRANDEDNESS: single

253 (D) TOPOLOGY: linear

## 255 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

257	CGGCAGCCAG CCTATTCTTT GGCCGGGTCG GTGCGAGTGG TCGGCTGGGC AGAGTGCACG	60
259	CTGCTTGGCG CCGCAGGTGA TCCCCCGGTC CACTCCCCGGG AGCAGTGATG TTGGGCAACT	120
261	CTGCGCCGGG GCCTGCGACC CGCGAGGCGG GCTCGGCGCT GCTAGCATTG CAGCAGACGG	180
263	CGCTCCAAGA GGACCAGGAG AATATCAACC CGGAAAAGGC AGCGCCCGTC CACAACCGC	240
265	GGACCCGGGC CGCGCTGGCG GTACTGAAGT CCGGGAACCC GCGGGGTCTA GCGCAGCAGC	300
267	AGAGGCCGAA GACGAGACGG GTTGACCCCC TTAAGGATCT TCCTGTAAAT GATGAGCATG	360
269	TCACCGTTCC TCCTTGGAAG GCAAACAGTA AACAGCCTGC GTTCACCATT CATGTGGATG	420
271	AAGCAGAAAA AGAAGCTCAG AAGAAGCCAG CTGAATCTCA AAAAATAGAG CGTGAAGATG	480
273	CCCTGGCTTT TAATTCAGCC ATTAGTTTAC CTGGACCCAG AAAACCATTG GTCCCTCTTG	540
275	ATTATCCAAT GGATGGTAGT TTTGAGTCAC CACATACTAT GGACATGTCA ATTGTATTAG	600
277	AAGATGAAAA GCCAGTGAGT GTTAATGAAG TACCAGACTA CCATGAGGAT ATTCACACAT	660
279	ACCTTAGGGA AATGGAGGTT AAATGTAAAC CTAAAGTGGG TTACATGAAG AAACAGCCAG	720
281	ACATCACTAA CAGTATGAGA GCTATCCTCG TGGACTGGTT AGTTGAAGTA GGAGAAGAAT	780
283	ATAAACTACA GAATGAGACC CTGCATTTGG CTGTGAACTA CATTGATAGG TTCCTGTCTT	840
285	CCATGTCAGT GCTGAGAGGA AAACCTTCAGC TTGTGGGCAC TGCTGCTATG CTGTTAGCCT	900
287	CAAAGTTTGA AGAAATATAC CCCCCAGAAG TAGCAGAGTT TGTGTACATT ACAGATGATA	960
289	CCTACACCAA GAAACAAGTT CTGAGAATGG AGCATCTAGT TTTGAAAGTC CTTACTTTTG	1020
291	ACTTAGCTGC TCCAACAGTA AATCAGTTTC TTACCCAATA CTTTCTGCAT CAGCAGCCTG	1080
293	CAAACCTGCA AGTTGAAAGT TTAGCAATGT TTTTGGGAGA ATTAAGTTTG ATAGATGCTG	1140
295	ACCCATACCT CAAGTATTTG CCATCAGTTA TTGCTGGAGC TGCCTTTCAT TTAGCACTCT	1200
297	ACACAGTCAC GGGACAAAGC TGGCCTGAAT CATTAATACG AAAGACTGGA TATACCCTGG	1260
299	AAAGTCTTAA GCCTTGTCTC ATGGACCTTC ACCAGACCTA CCTCAAAGCA CCACAGCATG	1320
301	CACAACAGTC AATAAGAGAA AAGTACAAAA ATTCAAAGTA TCATGGTGTG TCTCTCCTCA	1380
303	ACCCACCAGA GACACTAAAT CTGTAACAAT GAAAGACTGC CTTTGTTTTT TAAGATGTAA	1440
305	ATCACTCAAA GTATATGGTG TACAGTTTTT AACTTAGGTT TTTAATTTTA CAATCATTTT	1500
307	TGAATACAGA AGTTGTGGCC AAGTACAAAT TATGGTATCT ATTACTTTTT AAATGGTTTT	1560
309	AATTTGTATA TCTTTTGTAT ATGTATCTGT CTTAGATATT TGGCTAATTT TAAGTGGTTT	1620
311	TGTTAAAGTA TTAATGATGC CAGCTGCCG	1649

## 314 (2) INFORMATION FOR SEQ ID NO: 9:

## 316 (i) SEQUENCE CHARACTERISTICS:

317 (A) LENGTH: 175 base pairs

318 (B) TYPE: nucleic acid

319 (C) STRANDEDNESS: single

320 (D) TOPOLOGY: linear

## 322 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

324	ACCCACTCGT GAGTCCAACG GTCTTTTCTG CAGAAAGGAG GACTTTCCTT TCAGGGGTCT	60
326	TTCTGGGGCT CTTACTATAA AAGGGGACCA ACTCTCCCTT TGTCATATCT TGTTTCTGAT	120
328	GACAAAAAAT AACACATTGT TAAAATTGTA AAATTTAAAC ATGAAATATA AATTA	175

## 331 (2) INFORMATION FOR SEQ ID NO: 10:

## 333 (i) SEQUENCE CHARACTERISTICS:

334 (A) LENGTH: 166 base pairs

335 (B) TYPE: nucleic acid

## RAW SEQUENCE LISTING

DATE: 12/18/2001

PATENT APPLICATION: US/09/974,546

TIME: 16:40:03

Input Set : N:\Crif3\RULE60\09974546.raw

Output Set : N:\CRF3\12182001\I974546.raw

```

336         (C) STRANDEDNESS: single
337         (D) TOPOLOGY: linear
339         (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:
341 GTTTCGCTCC ACATTCATCC TTTCTTACTG GGCCTGATG TTGAGAGCAT CAGGCAGGGT      60
343 ATAATGTTAT GTTGACGTAA CAAACACCCT CAATATCTCA GTGGCTTAAA ATGACAACGA      120
345 TCTTTTTTTT GTTGTTTGT TTATGCTCTA TATCACCCAG GGATCA      166
348 (2) INFORMATION FOR SEQ ID NO: 11:
350     (i) SEQUENCE CHARACTERISTICS:
351         (A) LENGTH: 107 base pairs
352         (B) TYPE: nucleic acid
353         (C) STRANDEDNESS: single
354         (D) TOPOLOGY: linear
356     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:
358 TGCTCTGCCC CACATCTGAA CAAGCTAATA AGAAAGCCCG ATGTTCTTTC CTTTGGTGCC      60
360 ATTGGGAAAT TCAAACCATG CACAACCTG CCTGTATGAA GGGCGCA      107
363 (2) INFORMATION FOR SEQ ID NO: 12:
365     (i) SEQUENCE CHARACTERISTICS:
366         (A) LENGTH: 183 base pairs
367         (B) TYPE: nucleic acid
368         (C) STRANDEDNESS: single
369         (D) TOPOLOGY: linear
371     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:
373 CAACCTTAGC CCCTCTCCTC TTCTTCACGA TGCCATTCTG CCATTTCTGT TTTGTGGTAG      60
375 ACAGGTTGGC CCAGGCACTC TAAGGCCAG GCTGGCACAG GTTGGCCCAG GCACTTCAAG      120
377 CCTAAGTCCA TTTACAGTTT CTATTCCATC TCTTCCTAAA GAAGAGGAGA GGGGCTAAG      180
379 TTG      183
382 (2) INFORMATION FOR SEQ ID NO: 13:
384     (i) SEQUENCE CHARACTERISTICS:
385         (A) LENGTH: 92 base pairs
386         (B) TYPE: nucleic acid
387         (C) STRANDEDNESS: single
388         (D) TOPOLOGY: linear
390     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:
392 AAACAAACGT CTTTGGGTAA AATTCTATTT CTTTAAATGT TTTAAATAT TTGTAGTCAC      60
394 TAATTGTAAG TCATATTCCT CTTTGTCCAG CT      92
397 (2) INFORMATION FOR SEQ ID NO: 14:
399     (i) SEQUENCE CHARACTERISTICS:
400         (A) LENGTH: 182 base pairs
401         (B) TYPE: nucleic acid
402         (C) STRANDEDNESS: single
403         (D) TOPOLOGY: linear
405     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:
407 GATGTAATTA AAGCTGTAGA TGAGGGCTAT CGACTGCCAC CCCCCATGGA CTGCCCAGCT      60
409 GCCTTGATC AGCTGATGCT GGACTGCTGG CAGAAAGACA GGAACAACAG ACCCAAGTTT      120
411 GAGCAGATTG TTAGTATTCT GGACAAGCTT ATCCGGAATC CCGGCAGCCT GAAGGATCAT      180
413 CA      182
416 (2) INFORMATION FOR SEQ ID NO: 15:
418     (i) SEQUENCE CHARACTERISTICS:
419         (A) LENGTH: 174 base pairs

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/974,546

DATE: 12/18/2001

TIME: 16:40:04

Input Set : N:\Crf3\RULE60\09974546.raw

Output Set: N:\CRF3\12182001\I974546.raw

L:30 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]

L:31 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]